

1/22

Figure 1

1 GAATTCGCGGCCGCGCTCGACAGATGCCTTCTTCTGCCTGAGATTACACCCCACTAGCCAA 60
61 CCACTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCAGGGCAGCTCACCTGG 120
121 ACCCGACTGCCCCAAGGGTTCAAGAACTCCCCGACCATCTTGACGAAGCCCTACACAGG 180
181 GACCTGGCCAACTTCAGGATCCAACACCCTCAGGTGACCCTCCTCCAGTACGTGGATGAC 240
241 CTGCTTCTGGCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGGCACTACTGCTG 300
301 GAATTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTTGCAGGAGAGAG 360
361 GTAACATACTTGGGGTACAGTTTGCGGGGCGGGCAGCGATGGCTGACGGAGGCACGGAAG 420
421 AAAACTGTAGTCCAGATACCGGCCCCAACCACAGCCAAACAAGTGAGAGAGTTTTTGGGG 480
481 ACAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGCAGCCTTAGCAGCCCCACTCTAC 540
541 CCGCTAACCAAAGAAAAAGGGGATTCTCCTGGGCTCCTGAGCACCAGAAGGCATTTGAT 600
601 GCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACATAACCC 660
661 TTTACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTTAACCACAAACCTA 720
721 GGACCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTGTAGCCAGTGGT 780
781 TGGCCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGGACGCTGACAAA 840
841 TTGACTTTGGGACAGAATATAACTGTAATAGCCCCCATGCATTGGAGAACATCGTTCGG 900
901 CAGCCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAAGCCTGCTTCTC 960
961 ACAGAGAGGGTCACTTTCGCTCCACCAGCCGCTCTCAACCCTGCCACTCTTCTGCCTGAA 1020
1021 GAGACTGATGAACCAGTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTC 1080
1081 CGCAAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGTTCACTGACGGA 1140
1141 AGCAGCTATGTGGTGGAAGGTAAGAGGATGGCTGGGGCGGCAGTGGTGGACGGGACCCGC 1200
1201 ACGATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAAGGCTGAGCTCATGGCC 1260

2/22

Figure 1 cont.

1261 CTCACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATACGGACAGCAGG 1320
1321 TATGCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGGGGTTGCTTACC 1380
1381 TCAGCAGGGAGGGAAATAAAGAACAAGAGGAAATTCTAAGCCTATTAGAAGCCTTACAT 1440
1441 TTGCCAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCAAAGATCTCATA 1500
1501 TCTAGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGGCTGTTAACCTT 1560
1561 CTGCCTATAATAGAAACGCCCAAAGCCCCAGAACCCAGACGACAGTACACCCTAGAAGAC 1620
1621 TGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGACCTGCTATACC 1680
1681 TCATATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCCAACAGATACAT 1740
1741 CGTCTAACCACCTAGGAACTAAACACCTGCAGCAGTTGGTCAGAACATCCCCTTATCAT 1800
1801 GTTCTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGCCCTGCCAGCTG 1860
1861 GTTAATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAAGCCACCCAGGC 1920
1921 GCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAACAAATATCTA 1980
1981 TTGGTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACT 2040
2041 TCAACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTTCCAAGATTTGGAATACCTAAG 2100
2101 GTAATAGGGTCAGACAATGGTCCAGCTTTCGTTGCCCAGGTAAGTCAGGGACTGGCCAAG 2160
2161 ATATTGGGGATTGATTGGAAACTGCATTGTGCATACAGACCCCAAAGCTCAGGACAGGTA 2220
2221 GAGAGGATGAATAGAACCATTAAAGAGACCCTTACCAAATTGACCACAGAGACTGGCATT 2280
2281 AATGATTGGATGGCTCTCCTGCCCTTTGTGCTTTTTAGGGTGAGGAACACCCCTGGACAG 2340
2341 TTTGGGCTGACCCCTATGAATTGCTCTACGGGGGACCCCCCGTTGGCAGAAATTGCC 2400
2401 TTTGCACATAGTGCTGATGTGCTGCTTTCCAGCCTTTGTTCTCTAGGCTCAAGGCGCTC 2460
2461 GAGTGGGTGAGGCAGCGAGCGTGGAAGCAGCTCCGGGAGGCCTACTCAGGAGGAGACTTG 2520

Figure 1 cont.

2521 CAAGTTCACATCGCTTCCAAGTTGGAGATTCAGTCTATGTTAGACGCCACCGTGCAGGA 2580
2581 AACCTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACACCAACGGCTGTG 2640
2641 AAAGTCGAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCCGGCGCCACCTCCC 2700
2701 GATTCGGGGTGGAAAGCCGAAAAGACTGAAAATCCCCTTAAGCTTCGCCTCCATCGCGTG 2760
2761 GTTCCTTACTCTGTCAATAACTCCTCAAGTTAATGGTAAACGCCTTGTGGACAGCCCGAA 2820
2821 CTCCCATAAACCCCTTATCTCTCACCTGGTTACTTACTGACTCCGGTACAGGTATTAATAT 2880
2881 TAACAGCACTCAAGGGGAGGCTCCCTTGGGGACCTGGTGGCCTGAATTATATGTCTGCCT 2940
2941 TCGATCAGTAATCCCTGGTCTCAATGACCAGGCCACACCCCCGATGTACTCCGTGCTTA 3000
3001 CGGGTTTTACGTTTGCCCGAGGACCCCCAAATAATGAAGAATATTGTGGAAATCCTCAGGA 3060
3061 TTTCCCTTTGCAAGCAATGGAGCTGCATAACTTCTAATGATGGGAATTGGAAATGGCCAGT 3120
3121 CTCTCAGCAAGACAGAGTAAGTTACTCTTTTGTAAACAATCCTACCAGTTATAATCAATT 3180
3181 TAATTATGGCCATGGGAGATGGAAAGATTGGCAACAGCGGGTACAAAAAGATGTACGAAA 3240
3241 TAAGCAAATAAGCTGTCATTGTTAGACCTAGATTACTTAAAAATAAGTTTCACTAAAAA 3300
3301 AAAAAAAAAAAAAAAAAAAAAA 3320

4/22

Figure 2

1 TGTGGGCCCCAGCGCGCTTGGAATAAAAAATCCTCTTGCTGTTTGCATCAAGACCGCTTCT 60

61 CGTGAGTGATTTGGGGTGTGCGCTCTTCCGAGCCCGGACGAGGGGGATTGTTCTTTTACT 120

121 GGCCTTTCATTGGTGCGTTGGCCGGGAAATCCTGCGACCACCCCTTACACCCGAGAACC 180

181 GACTTGAGGTAAAGGGATCCCCTTTGGAACGTGTGTGTGTGTCGGCCGGCGTCTCTGTT 240

241 CTGAGTGTCTGTTTTCGGTGATGCGCGCTTTCGGTTTGCAGCTGTCCTCTCAGACCGTAA 300

301 GGA CTGGAGGACTGTGATCAGCAGACGTGCTAGGAGGATCACAGGCTGCCACCCCTGGGGG 360

361 ACGCCCCGGGAGGTGGGGAGAGCCAGGGACGCCTGGTGGTCTCCTACTGTGGTTCAGAGG 420

421 ACCGAGTTCTGTTGTTGAAGCGAAAGCTTCCCCCTCCGCGGCCGTCCGACTCTTTTGCCT 480

481 GCTTGTGGAAGACGCGGACGGGTGCGGTGTGTCTGGATCTGTTGGTTTCTGTCTCGTGTG 540

541 TCTTTGTCTTGTGCGTCTTGTCTACAGTTTTAATATGGGACAGACAGTGACTACCCCCC 600

601 TTAGTTTGACTCTCGACCATTGGA CTGAAGTTAGATCCAGGGCTCATAATTTGTCAGTTC 660

661 AGGTTAAGAAGGGACCTTGGCAGACTTTCTGTGCCTCTGAATGGCCAACATTGATGTTG 720

721 GATGGCCATCAGAGGGGACCTTTAATTCTGAAATTATCCTGGCTGTTAAGGCAATCATT 780

781 TTCAGACTGGACCCGGCTCTCATCCTGATCAGGAGCCCTATATCCTTACGTGGCAAGATT 840

841 TGGCAGAAGATCCTCCGCCATGGGTTAAACCATGGCTAAATAAACCAAGAAAGCCAGGTC 900

901 CCCGAATCCTGGCTCTTGGAGAGAAAAACAAACTCGGCCGAAAAAGTCGAGCCCTCTT 960

961 CCTCGTATCTACCCCGAGATCGAGGAGCCGCCGACTTGGCCGGAACCCCAACCTGTTCCC 1020

1021 CCACCCCTTATCCAGCACAGGGTGCTGTGAGGGGACCTCTGCCCCCTCCTGGAGCTCCGG 1080

1081 TGGTGGAGGGACCTGCTGCCGGGACTCGGAGCCGGAGAGGCGCCACCCCGGAGCGGACAG 1140

1141 ACGAGATCGCGATATTACCGCTGCGCACCTATGGCCCTCCCATGCCAGGGGGCCAATTGC 1200

1201 AGCCCCCTCCAGTATTGGCCCTTTTCTTCTGCAGATCTCTATAATTGGAAAACTAACCATC 1260

5/22

Figure 2 cont.

1261 CCCCTTTCTCGGAGGATCCCAACGCCTCACGGGGTTGGTGGAGTCCCTTATGTTCTCTC 1320
1321 ACCAGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTCACAACCGAGGAGC 1380
1381 GAGAGAGAATTCTGTTAGAGGCTAGAAAAATGTTCTGGGGCCGACGGGCGACCCACGC 1440
1441 AGTTGCAAAATGAGATTGACATGGGATTTCCCTTGACTCGCCCCGGTTGGGACTACAACA 1500
1501 CGGCTGAAGGTAGGGAGAGCTTGAAAATCTATCGCCAGGCTCTGGTGGCGGGTCTCCGGG 1560
1561 GCGCCTCAAGACGGCCCACTAATTTGGCTAAGGTAAGAGAGGTGATGCAGGGACCGAACG 1620
1621 AACCTCCCTCGGTATTTCTTGAGAGGCTCATGGAAGCCTTCAGGCGGTTACCCCTTTTG 1680
1681 ATCCTACCTCAGAGGCCCAGAAAGCCTCAGTGGCCCTGGCCTTCATTGGGCAGTCGGCTC 1740
1741 TGGATATCAGGAAGAACTTCAGAGACTGGAAGGGTTACAGGAGGCTGAGTTACGTGATC 1800
1801 TAGTGAGAGAGGCAGAGAAGGTGTATTACAGAAGGGAGACAGAAGAGGAGAAGGAACAGA 1860
1861 GAAAAGAAAAGGAGAGAGAAGAAAGGGAGGAAAGACGTGATAGACGGCAAGAGAAGAATT 1920
1921 TGAATAAGATCTTGGCCGCAGTGGTTGAAGGGAAGAGCAGCAGGGAGAGAGAGAGATT 1980
1981 TTAGGAAAAATTAGGTGAGGCCCTAGACAGTCAGGGAACCTGGGCAATAGGACCCCACTCG 2040
2041 ACAAGGACCAGTGTGCGTATTGTAAAGAAAAGGACACTGGGCAAGGAAGTCCCCAAGA 2100
2101 AGGGAAACAAAGGACCGAAGTCCTAGCTCTAGAAGAAGATAAAGATTAGGGGAGACGGGT 2160
2161 TCGGACCCCTCCCCGAGCCAGGGTAACTTTGAAGGTGGAGGGGCAACCAGTTGAGTTC 2220
2221 CTGGTTGATACCGGAGCGGAGCATTGAGTGCTGCTACAACCATTAGGAAAATAAAAGAA 2280
2281 AAAAAATCCTGGGTGATGGGTGCCACAGGGCAACGGCAGTATCCATGGACTACCCGAAGA 2340
2341 ACCGTTGACTTGGGAGTGGGACGGGTAACCCACTCGTTTCTGGTCATCCCTGAGTGCCCA 2400
2401 GTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAATTTCTTTTGAACAA 2460
2461 GGAAGACCAGAAGTGTCTGTGAATAACAAACCCATCACTGTGTTGACCCTCCAATTAGAT 2520

6/22

Figure 2 cont.

2521 GATGAATATCGACTATATTCTCCCCAAGTAAAGCCTGATCAAGATATACAGTCCTGGTTG 2580

2581 GAGCAGTTTCCCCAAGCCTGGGCAGAAACCGCAGGGATGGGTTTGGCAAAGCAAGTTCCC 2640

2641 CCACAGGTTATTCAACTGAAGGCCAGTGCTACACCAGTATCAGTCAGACAGTACCCCTTG 2700

2701 AGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAAGATTAATCCAACAGGGCATC 2760

2761 CTAGTTCCTGTCCAATCCCCTTGGAATACTCCCCTGCTACCGTTAGGAAGCCTGGGACC 2820

2821 AATGATTATCGACCAGTACAGGACTTGAGAGAGGTCAATAAAAGGGTGCAGGACATACAC 2880

2881 CCAACGGTCCCGAACCCTTATAACCTCTTGAGCGCCCTCCCGCCTGAACGGAAGTGGTAC 2940

2941 ACAGTATTGGACTTAAAAGATGCCTTCTTCTGCCTGAGATTACACCCCACTAGCCAACCA 3000

3001 CTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCAGGAGCTCACCTGGACC 3060

3061 CGACTGCCCCAAGGGTTCAAGAAGTCCCCGACCATCTTTGACGAAGCCCTACACAGGGAC 3120

3121 CTGGCCAACCTTCAGGATCCAACACCCTCAGGTGACCCTCCTCCAGTACGTGGATGACCTG 3180

3181 CTTCTGBCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGGCACTACTGCTGGAA 3240

3241 TTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCCAGATTTGCAGGAGAGAGGTA 3300

3301 ACATACTTGGGGTACAGTTTGCAGGGGCGGGCAGCGATGGCTGACGGAGGCACGGAAGAAA 3360

3361 ACTGTAGTCCAGATACCGGCCCCAACCACAGCCAAACAAGTGAGAGAGTTTTTGGGGACA 3420

3421 GCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGCAGCCTTAGCAGCCCCACTCTACCCG 3480

3481 CTAACCAAAGAAAAAGGGGATTCTCCTGGGCTCCTGAGCACCAGAAGGCATTTGATGCT 3540

3541 ATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACCTAAACCCTTT 3600

3601 ACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCCGAGGAGTTTAAACCCAAACCCTAGGA 3660

3661 CCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTGTAGCCAGTGGTTGG 3720

3721 CCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGGACGCTGACAAATTG 3780

7/22

Figure 2 cont.

3781 ACTTTGGGACAGAATATAACTGTAATAGCCCCCATGCATTGGAGAACATCGTTCGGCAG 3840

3841 CCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAAGCCTGCTTCTCACA 3900

3901 GAGAGGGTCACTTTGCTCCACCAGCCGCTCTCAACCCTGCCACTCTTCTGCCTGAAGAG 3960

3961 ACTGATGAACCACTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTCCGC 4020

4021 AAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGTTCCTGACGGAAGC 4080

4081 AGCTATGTGGTGAAGGTAAGAGGATGGCTGGGGCGGCAGTGGTGGACGGGACCCGCACG 4140

4141 ATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAGGCTGAGCTCATGGCCCTC 4200

4201 ACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATACGGACAGCAGGTAT 4260

4261 GCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGGGGTTGCTTACCTCA 4320

4321 GCAGGGAGGGAAATAAAGAACAAGAGGAAATTCTAAGCCTATTAGAAGCCTTACATTTG 4380

4381 CCAAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCAAAGATCTCATATCT 4440

4441 AGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGGCTGTTAACCTTCTG 4500

4501 CCTATAATAGAAACGCCCAAAGCCCCAGAACCAGACGACAGTACACCCTAGAAGACTGG 4560

4561 CAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGACCTGCTATACCTCA 4620

4621 TATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCCAACAGATACATCGT 4680

4681 CTAACCCACCTAGGAATAAACACCTGCAGCAGTTGGTCAGAACATCCCCTTATCATGTT 4740

4741 CTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGCCCTGCCAGCTGGTT 4800

4801 AATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAAGCCACCCAGGCGCT 4860

4861 CACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAACAAATATCTATTG 4920

4921 GTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACTTCA 4980

4981 ACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTTCCAAGATTTGGAATACCTAAGGTA 5040

8/22

Figure 2 cont.

5041 ATAGGGTCAGACAATGGTCCAGCTTTCGTTGCCAGGTAAGTCAGGGACTGGCCAAGATA 5100
5101 TTGGGGATTGATTGGAACTGCATTGTGCATACAGACCCCAAAGCTCAGGACAGGTAGAG 5160
5161 AGGATGAATAGAACCATTAAAGAGACCCTTACCAAATTGACCACAGAGACTGGCATTAAAT 5220
5221 GATTGGATGGCTCTCCTGCCCTTTGTGCTTTTTAGGGTGAGGAACACCCCTGGACAGTTT 5280
5281 GGGCTGACCCCTATGAATTGCTCTACGGGGACCCCCCGTTGGCAGAAATTGCCTTT 5340
5341 GCACATAGTGCTGATGTGCTGCTTTCCAGCCTTTGTTCTCTAGGCTCAAGGCGCTCGAG 5400
5401 TGGGTGAGGCAGCGAGCGTGAAGCAGCTCCGGGAGGCCTACTCAGGAGGAGACTTGCAA 5460
5461 GTTCCACATCGCTTCCAAGTTGGAGATTCACTCTATGTTAGACGCCACCGTGCAGGAAAC 5520
5521 CTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACACCAACGGCTGTGAAA 5580
5581 GTCGAAGGAATCCCCACCTGGATCCATGCATCCACGTTAAGCYGGCGCCACCTCCCGAC 5640
5641 TCGGGGTGGAGAGCCGAAAAGACTGAGAATCCCCTTAAGCTTCGCCTCCATCGCCTGGTT 5700
5701 CCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATAGACAGCTCGAACCC 5760
5761 CCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACGGGTGTCAGTGTAAA 5820
5821 TAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTGCATTTCTGCCTCCG 5880
5881 ATTGATTAACCCCGCTGTTAARAGCACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTA 5940
5941 TTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTCTGGGGAATCCTTCTGTAG 6000
6001 GAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCGATCTCTCTCCAGGA 6060
6061 CCGGGTAAAATTCTCCTTTGTCAATTCGGGCCGGGCAAGTACAAAATGATGAACTATA 6120
6121 TAAAGATAAGAGCTGCTCCCATCAGACTTAGATTATCTAAAGATAAGTTTCACTGAAAG 6180
6181 GAAAACAGGAAAATATTCAAAAGTGGATAAATGGTATGAGCTGGGGAATAGTTTTTTATT 6240
6241 ATATGGCGGGGGAGCAGGGTCCACTTTAACCATTGCGCTTAGGATAGAGACGGGGACAGA 6300

9/22

Figure 2 cont.

6301 ACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCCGGCCCTGGA 6360
6361 GCCACCGCATAAATTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATAACACAGCC 6420
6421 GCCTAGCAACAGTACCACTGGATTGATTCTTACCAACACGCCTAGAAACTCCCCAGGTGT 6480
6481 TCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTTCCAAGCCATCAA 6540
6541 CTCCACCGACCCTGATGCCACTTCTTCTTGTTGGCTTTGTCTATCCTCAGGGCCTCCTTA 6600
6601 TTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGAAATCAATG 6660
6661 TACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGGACATGCAAT 6720
6721 AGGAAAAGCTCCCCCATCCCACCAACACCTTTGCTATAGTACTGTGGTTTATGAGCAGGC 6780
6781 CTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAATACTGGGTT 6840
6841 AACCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTGATGGTCCA 6900
6901 AATCGTCCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGATGAATATGACTATCG 6960
6961 GTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTCGGATTAGG 7020
7021 GACGGCCGTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGACCACAGCAGCTAGA 7080
7081 GAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGAGCCTTAAAGGAGTC 7140
7141 TGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTTCTACAGAACCGGAG 7200
7201 GGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTATGTGCAGCCTTAAAAGAAGAATG 7260
7261 TTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTTAGAAAAAA 7320
7321 GTTAGAGAGGCCTCGAAGGGAAAGAGAGGCTGACCAGGGTGGTTTGAAGGATGGTTCAA 7380
7381 CAGGTCTCCTTGATGACCACCCTGCTTTCTGCTCTGACGGGGCCCCTAGTAGTCCTGCT 7440
7441 CCTGTTACTTACAGTTGGGCCTTGCTTAATTAATAGGTTTGTGCTTTGTTAGAGAACG 7500
7501 AGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCTTCTGAGCCAAGG 7560

10/22

Figure 2 cont.

7561	AGAAACTGACCTCTAGCCTTCCCAGTTCTAAGATTAGAAGCTATTAACAAGACAAGAAGTG	7620
7621	GGGAATGAAAGGATGAAAAATGCAACCTAACCTCCCAGAAGCCAGGAAGTTAATAAAAAG	7680
7681	CTCTAAATGCCCCGAATTMCAGACCCTGCTGGCTGCCAGTAAATAGGTAGAAGGTCACA	7740
7741	CTTCCTATTGTTCCAGGGCCTGCTATCCTGGCCTAAGTAAGATAACAGGAAATGAGTTGA	7800
7801	CTAATCGCTTATCTGGATTCTGTAAAGCTGACTGGCACCATAGAAGAATTGATTACACAT	7860
7861	TGACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCAGGAGCCCACGCAG	7920
7921	ATGCGGACCTCCGGAGCTATTTTAAAGTATTGGTCCACGGAGCGCGGGCTCTCGATATT	7980
7981	TTAAAGTATTGGTCCATGGAGCGCGGGCTCTCGATATTTTAAAGTATTGGTTTGTGAC	8040
8041	GCACAGGCTTTGTTGTGAACCCCATAAAGCTGTCCCGATTCCGCACTCGGGGCCGCAGT	8100
8101	CCTCTACCCCTGCGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGAATAAAAATCCTCT	8160
8161	TGCTGTTTGCATCAAAAAAAAAAAAAAAAAAAAAA	8196

[illegible]

Figure 3

11/22

1 GTGGTGTACGACTGTGGGCCCCAGCGCGCTTGAATAAAAAATCCTCTTGCTGTTTGCATC 60

61 AAGACCGCTTCTCGTGAGTGATTTGGGGTGTGCGCTCTTCCGAGCCCGGACGAGGGGGAT 120

121 TGTTCTTTTACTGGCCTTTTCATTTGGTGCGTTGGCCGGGAAATCCTGCGACCACCCCTTA 180

181 CACCCGAGAACC GACTTGGAGGTAAAGGGATCCCCCTTTGGAACGTGTGTGTGTGTCGGCC 240

241 GCGCTCTCTGTTCTGAGTGTCTGTTTTCGGTGATGCGCGCTTTCGGTTTGCAGCTGTCTCT 300

301 CTCAGACCGTAAGGACTGGAGGACTGTGATCAGCAGACGTGCTAGGAGGATCACAGGCTG 360

361 CCACCCTGGGGGACGCCCCGGGAGGTGGGGAGAGCCAGGGACGCCTGGTGGTCTCCTACT 420

421 GTCGGTCAGAGGACCGAGTTCTGTTGTTGAAGCGAAAGCTTCCCCCTCCGCGGCCGTCGG 480

481 ACTCTTTTGCCTGCTTGTGGAAGACGCGGACGGGTGCGGTGTGTCTGGATCTGTTGGTTT 540

541 CTGTCTCGTGTGTCTTTGTCTTGTGCGTCCTTGTCTACAGTTTTTAATATGGGACAGACAG 600
MetGlyGlnThrV

601 TGACTACCCCCCTTAGTTTGA CTCTCGACCATTGGACTGAAGTTAGATCCAGGGCTCATA 660
alThrThrProLeuSerLeuThrLeuAspHisTrpThrGluValArgSerArgAlaHisA

661 ATTTGTCTAGTTTCAGGTTAAGAAGGGACCTTGGCAGACTTTCTGTGCCTCTGAATGGCCAA 720
snLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysAlaSerGluTrpProT

721 CATTTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATTCTGAAATTATCCTGGCTGTTA 780
hrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValL

781 AGGCAATCATTTTTTCAGACTGGACCCGGCTCTCATCCTGATCAGGAGCCCTATATCCTTA 840
ysAlaIleIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleLeuT

841 CGTGGCAAGATTTGGCAGAAGATCCTCCGCCATGGGTAAACCATGGCTAAATAAACCAA 900
hrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrpLeuAsnLysProA

901 GAAAGCCAGGTCCCCGAATCCTGGCTCTTGGAGAGAAAAACAAACACTCGGCCGAAAAAG 960
rgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHisSerAlaGluLysV

961 TCGAGCCCTCTTCCTCGTATCTACCCCGAGATCGAGGAGCCCGGACTTGGCCGGAACCC 1020
alGluProSerSerSerTyrLeuProArgAspArgGlyAlaAlaAspLeuAlaGlyThrP

1021 CAACCTGTTCCCCACCCCTTATCCAGCACAGGGTGCTGTGAGGGGACCTCTGCCCTC 1080
roThrCysSerProThrProLeuSerSerThrGlyCysCysGluGlyThrSerAlaProP

12/22

Figure 3 cont.

1081	CTGGAGCTCCGGTGGTGGAGGGACCTGCTGCCGGGACTCGGAGCCGGAGAGGCGCCACCC roGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAlaThrP	1140
1141	CGGAGCGGACAGACGAGATCGCGATATTACCGCTGCGCACCTATGGCCCTCCCATGCCAG roGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGlyProProMetProG	1200
1201	GGGGCCAATTGCAGCCCCTCCAGTATTGGCCCTTTTCTTCTGCAGATCTCTATAATTGGA lyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAspLeuTyrAsnTrpL	1260
1261	AAACTAACCATCCCCCTTTCTCGGAGGATCCCCAACGCCTCACGGGGTTGGTGGAGTCCC ysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGlyLeuValGluSerL	1320
1321	TTATGTTCTCTCACCAGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTCA euMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheT	1380
1381	CAACCGAGGAGCGAGAGAGAATTCTGTTAGAGGCTAGAAAAAATGTTCTTGGGGCCGACG hrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAlaAspG	1440
1441	GGCGACCCACGCAGTTGCAAAATGAGATTGACATGGGATTTCCCTTGACTCGCCCCGGTT lyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeuThrArgProGlyT	1500
1501	GGGACTACAACACGGCTGAAGGTAGGGAGAGCTTGAAAATCTATCGCCAGGCTCTGGTGG rpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArgGlnAlaLeuValA	1560
1561	CGGGTCTCCGGGGCGCCTCAAGACGGCCCACTAATTTGGCTAAGGTAAGAGAGGTGATGC laGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysValArgGluValMetG	1620
1621	AGGGACCGAACGAACCTCCCTCGGTATTTCTTGAGAGGCTCATGGAAGCCTTCAGGCGGT lnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGluAlaPheArgArgP	1680
1681	TCACCCCTTTTGATCCTACCTCAGAGGCCAGAAAAGCCTCAGTGGCCCTGGCCTTCATTG heThrProPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPheIleG	1740
1741	GGCAGTCGGCTCTGGATATCAGGAAGAACTTCAGAGACTGGAAGGGTTACAGGAGGCTG lyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGlyLeuGlnGluAlaG	1800
1801	AGTTACGTGATCTAGTGAGAGAGGCAGAGAAGGTGTATTACAGAAGGGAGACAGAAGAGG luLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArgGluThrGluGluG	1860
1861	AGAAGGAACAGAGAAAAGAAAAGGAGAGAGAAGAAAGGGAGGAAAGACGTGATAGACGGC luLysGluGlnArgLysGluLysGluArgGluGluArgGluGluArgArgAspArgArgG	1920
1921	AAGAGAAGAATTTGACTAAGATCTTGGCCGCAGTGGTTGAAGGGAAGAGCAGCAGGGAGA lnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLysSerSerArgGluA	1980
1981	GAGAGAGAGATTTTAGGAAAATTAGGTCAGGCCCTAGACAGTCAGGGAACCTGGGCAATA rgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGlyAsnA	2040

Figure 3 cont

2041	GGACCCCACTCGACAAGGACCAGTGTGCGTATTGTAAAGAAAAAGGACACTGGGCAAGGA rgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGlyHisTrpAlaArgA	2100
2101	ACTGCCCCAAGAAGGGAAACAAAGGACCGAAGgTCCTAGCTCTAGAAGAAGATAAAGATT snCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGluGluAspLysAspE	2160
2161	AGGGGAGACGGGgTTCGGACCCCCTCCCCGAGCCCAGGGTAACTTTGAAGGTGGAGGGGC ndGlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysValGluGlyG	2220
2221	AACCAGTTGAGTTCCTGGTTGATACCGGAGCGGAGCATTCACTGCTGCTACAACCATTAG lnProValGluPheLeuValAspThrGlyAlaGluHisSerValLeuLeuGlnProLeuG	2280
2281	GAAAACTAAAAGAAAAAAATCCTGGGTGATGGGTGCCACAGGGCAACGGCAGTATCCAT lyLysLeuLysGluLysLysSerTrpValMetGlyAlaThrGlyGlnArgGlnTyrProT	2340
2341	GGACTACCCGAAGAACCgTTGACTTGGGAGTGGGACGGGTAAACCCACTCGTTTCTGGTCA rpThrThrArgArgThrValAspLeuGlyValGlyArgValThrHisSerPheLeuValI	2400
2401	TCCCTGAGTGCCcAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAA leProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnI	2460
2461	TTTCTTTTGAACAAGGAAGACCAGAAGTGTCTGTGAATAACAAACCCATCACTGTGTTGA leSerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleThrValLeuT	2520
2521	CCCTCCAATTAGATGATGAATATCGACTATATTCTCCCCAAGTAAAGCCTGATCAAGATA hrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnValLysProAspGlnAspI	2580
2581	TACAGTCCTGGTTGGAGCAGTTTCCCCAAGCCTGGGCAGAAACCGCAGGGATGGGTTTGG leGlnSerTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAlaGlyMetGlyLeuA	2640
2641	CAAAGCAAGTTCCCCCACAGGTTATTCAACTGAAGGCCAGTGCTACACCAGTATCAGTCA laLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThrProValSerValA	2700
2701	GACAGTACCCCTTGAGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAAGATTAA rgGlnTyrProLeuSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuI	2760
2761	TCCAACAGGGCATCCTAGTTCCTGTCCAATCCCCTTGAATACTCCCCTGCTACCGGTTA leGlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValA	2820
2821	GGAAGCCTGGGACCAATGATTATCGACCAGTACAGGACTTGAGAGAGGTCAATAAAAGGG rgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLysArgV	2880
2881	TGCAGGACATACACCCAACGGTCCCGAACCCCTTATAACCTCTTGAGCGCCCTCCCGCCTG alGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSerAlaLeuProProG	2940
2941	AACGGAACTGGTACACAGTATTGGACTTAAAAGATGCCTTCTTCTGCCTGAGATTACACC luArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisP	3000

14/22

Figure 3 cont.

3001	CCACTAGCCAACCACTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACC3GGC roThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyG	3060
3061	AGCTCACCTGGACCCGACTGCCCCAAGGGTTCAAGAACTCCCGACCATCTTTGACGAAG lnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluA	3120
3121	CCCTACACAGGGACCTGGCCAACTTCAGGATCCAACACCCTCAGGTGACCCTCCTCCAGT laLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnT	3180
3181	ACGTGGATGACCTGCTTCTGGCGGGAGCCACCAACAGGACTGCTTAGAAGGTACGAAGG yrValAspAspLeuLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysA	3240
3241	CACTACTGCTGGAATTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTT laLeuLeuLeuGluLeuSerAspLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleC	3300
3301	GCAGGAGAGAGGTAACATACTTGGGGTACAGTTTGGGGGGCGGGCAGCGATGGCTGACGG ysArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrG	3360
3361	AGGCACGGAAGAAAAGTGTAGTCCAGATACCGGCCCAACCACAGCCAAACAAGTGAGAG luAlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGlnValArgG	3420
3421	AGTTTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGGGACCTTAGCAG luPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaA	3480
3481	CCCCACTCTACCCGCTAACCAGAAAGAAAAAGGGGGATTCTCCTGGGCTCCTGAGCACCAGA laProLeuTyrProLeuThrLysGluLysGlyGlyPheSerTrpAlaProGluHisGlnL	3540
3541	AGGCATTTGATGCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACG ysAlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspV	3600
3601	TAACTAAACCCTTTACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTAA alThrLysProPheThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuT	3660
3661	CCCAAACCCTAGGACCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTG hrGlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeuAspProV	3720
3721	TAGCCAGTGGTTGGCCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGG alAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaValAlaIleLeuValLysA	3780
3781	ACGCTGACAAATTGACTTTGGGACAGAATATACTGTAATAGCCCCCATGCATTGGGAGA spAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaProHisAlaLeuGluA	3840
3841	ACATCGTTTCGGCAGCCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAA snIleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTyrGlnS	3900
3901	GCCTGCTTCTCACAGAGAGGGTCACTTTTCGCTCCACCAGCCGCTCTCAACCCTGCCACTC erLeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrL	3960

Figure 3 cont

3961	TTCTGCCTGAAGAGACTGATGAACCACTGACTCATGATTGCCATCAACTATTGATTGAGG euLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLeuIleGluG	4020
4021	AGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGT luThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGluValLeuThrTrpP	4080
4081	TCACTGACGGAAGCAGCTATGTGGTGGAAGGTAAGAGGATGGCTGGGGCGGCAGTGGTGG heThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGlyAlaAlaValValA	4140
4141	ACGGGACCCGACGATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAGGCTG spGlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaG	4200
4201	AGCTCATGGCCCTCACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATA luLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrT	4260
4261	CGGACAGCAGGTATGCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGG hrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgG	4320
4321	GGTTGCTTACCTCAGCAGGGAGGGAAATAAAGAACAAGAGGAAATTCTAAGCCTATTAG lyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIleLeuSerLeuLeuG	4380
4381	AAGCCTTACATTTGCCAAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCA luAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGlyHisGlnLysAlaL	4440
4441	AAGATCTCATATCTAGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCAGG ysAspLeuIleSerArgGlyAsnGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnA	4500
4501	CTGTTAACCTTCTGCCTATAATAGAAAACGCCCAAAGCCCCAGAAGCCAGACGACAGTACA laValAsnLeuLeuProIleIleGluThrProLysAlaProGluProArgArgGlnTyrT	4560
4561	CCCTAGAAGACTGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGA hrLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyT	4620
4621	CCTGCTATACCTCATATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCC hrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGlyLeuGluTyrValG	4680
4681	AACAGATACATCGTCTAAGCCACCTAGGAACTAAACACCTGCAGCAGTTGGTCAGAACAT lnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGlnLeuValArgThrS	4740
4741	CCCCTTATCATGTTCTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGC erProTyrHisValLeuArgLeuProGlyValAlaAspSerValValLysHisCysValP	4800
4801	CCTGCCAGCTGGTTAATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAA roCysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlys	4860
4861	GCCACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAA erHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLysTyrGlyA	4920

Figure 3 cont.

4921	ACAAATATCTATTGGTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTA snLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyrProThrL	4980
4981	AGAAAGAGACTTCAACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTCCAAGATTTG ysLysGluThrSerThrValValAlaLysLysIleLeuGluGluIlePheProArgPheG	5040
5041	GAATACCTAAGGTAATAGGGTCAGACAATGGTCCAGCTTTTCGTTGCCCAGGTAAGTCAGG lyIleProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGlnG	5100
5101	GACTGGCCAAGATATTGGGGATTGATTGGAACTGCATTGTGCATACAGACCCCAAAGCT lyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerS	5160
5161	CAGGACAGGTAGAGAGGATGAATAGAACCATTAAAGAGACCCTTACCAAATTGACCACAG erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrThrG	5220
5221	AGACTGGCATTAAATGATTGGATGGCTCTCCTGCCCTTTGTGCTTTTATAGGGTGAGGAACA luThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgValArgAsnT	5280
5281	CCCCTGGACAGTTTGGGCTGACCCCTATGAATTGCTCTACGGGGGACCCCCCCTTGG hrProGlyGlnPheGlyLeuThrProTyrGluLeuLeuTyrGlyGlyProProProLeuA	5340
5341	CAGAAATTGCCTTTGCACATAGTGCTGATGTGCTGCTTTCCAGCCTTTGTTCTCTAGGC laGluIleAlaPheAlaHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgL	5400
5401	TCAAGGCGCTCGAGTGGGTGAGGCAGCGAGCGTGAAGCAGCTCCGGGAGGCCTACTCAG euLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerG	5460
5461	GAGGAGACTTGCAAGTTCACATCGCTTCCAAGTTGGAGATTCACTCTATGTTAGACGCC lyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrValArgArgH	5520
5521	ACCGTGCAGGAAACCTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACAC lsArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrLeuValLeuLeuThrThrP	5580
5581	CAACGGCTGTGAAAGTCAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCCGG roThrAlaValLysValGluGlyIleProThrTrpIleHisAlaSerHisValLysProA MetHisProThrLeuSerArg	5640
5641	CGCCACCTCCCGACTCGGGGTGGAGAGCCGAAAAGACTGAGAATCCCCTTAAGCTTCGCC laProProProAspSerGlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgL ArgHisLeuProThrArgGlyGlyGluProLysArgLeuArgIleProLeuSerPheAla	5700
5701	TCCATCGCCTGGTTCCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATA euHisArgLeuValProTyrSerAsnAsnAsnSerProGlyGlnEnd SerIleAlaTrpPheLeuThrLeuThrIleThrProGlnAlaSerSerLysArgLeuIle	5760
5761	GACAGCTCGAACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACG AspSerSerAsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspThr	5820

Figure 3 cont

17/22

5821 GGTGTCACCTGTAAATAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTG 5880
GlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpTrpProGluLeu

5881 CATTCTGCCTCCGATTGATTAACCCCGCTGTTAAAGCAGACCTCCCAACCTAGTCCGT 5940
HisPheCysLeuArgLeuIleAsnProAlaValLysSerThrProProAsnLeuValArg

5941 AGTTATGGGTTCTATTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGG 6000
SerTyrGlyPheTyrCysCysProGlyThrGluLysGluLysTyrCysGlyGlySerGly

6001 GAATCCTTCTGTAGGAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCG 6060
GluSerPheCysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro

6061 ATCTCTCTCCAGGACCGGGTAAAATTCTCCTTTGTCAATTCGGGCCCGGGCAAGTACAAA 6120
IleSerLeuGlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLys

6121 ATGATGAAACTATATAAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATA 6180
MetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspTyrLeuLysIle

6181 AGTTTCACTGAAAGGAAAACAGGAAAATATTCAAAGTGGATAAATGGTATGAGCTGGGG 6240
SerPheThrGluArgLysThrGlyLysTyrSerLysValAspLysTrpTyrGluLeuGly

6241 AATAGTTTTTTTATTATATGGCGGGGGAGCAGGGTCCACTTTAACCATTGCGCTTAGGATA 6300
AsnSerPheLeuLeuTyrGlyGlyGlyAlaGlySerThrLeuThrIleArgLeuArgIle

6301 GAGACGGGGACAGAACCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGG 6360
GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAlaGluGlnGly

6361 CCCCCGGCCCTGGAGCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCCT 6420
ProProAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgPro

6421 GACATAACACAGCCGCTAGCAACAGTACCACTGGATTGATTCCTACCAACACGCCTAGA 6480
AspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrAsnThrProArg

6481 AACTCCCCAGGTGTTCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCT 6540
AsnSerProGlyValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAla

6541 TTCCAAGCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTGGCTTTGTCTATCC 6600
PheGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer

6601 TCAGGGCCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAG 6660
SerGlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGlu

6661 CATAGAAATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGG 6720
HisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGly

6721 AAGGGGACATGCATAGGAAAAGCTCCCCCATCCACCAACACCTTTGCTATAGTACTGTG 6780
LysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrVal

Figure 3 cont.

18/22

6781	GTTTATGAGCAGGCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCA ValTyrGluGlnAlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAla	6840
6841	TGCAATACTGGGTAAACCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTTC CysAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe	6900
6901	TGTGTCATGGTCCAAATCGTCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGAT CysValMetValGlnIleValProArgValTyrTyrHisProGluGluValValLeuAsp	6960
6961	GAATATGACTATCGGTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTA GluTyrAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaVal	7020
7021	ATGCTCGGATTAGGGACGGCCGTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGA MetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGly	7080
7081	CCACAGCAGCTAGAGAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGA ProGlnGlnLeuGluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArg	7140
7141	GCCTTAAAGGAGTCTGTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTT AlaLeuLysGluSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal	7200
7201	CTACAGAACCGGAGGGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTTATGTGCAGCC LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAla	7260
7261	TTAAAAGAAGAATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAAC LeuLysGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAsn	7320
7321	AAGCTTAGAAAAAAGTTAGAGAGGCGTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTT LysLeuArgLysLysLeuGluArgArgArgArgGluArgGluAlaAspGlnGlyTrpPhe	7380
7381	GAAGGATGGTTCAACAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGGGCCC GluGlyTrpPheAsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyPro	7440
7441	CTAGTAGTCCTGCTCCTGTTACTTACAGTTGGGCCTTGCTTAATTAATAGGTTTGTGGCC LeuValValLeuLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla	7500
7501	TTTGTTAGAGAACGAGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGC PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGly	7560
7561	CTTCTGAGCCAAGGAGAACTGACCTCTAGCCTTCCCAGTTCTAAGATTAGAACTATTAA LeuLeuSerGlnGlyGluThrAspLeuEnd	7620
7621	CAAGACAAGAAGTGGGGAATGAAAGGATGAAATGCAACCTAACCCCTCCCAGAACCCAGG	7680
7681	AAGTTAATAAAAAGCTCTAAATGCCCCCGAATTACAGACCCTGCTGGCTGCCAGTAAATA	7740

Figure 3 cont.

19/22

7741 GGTAGAAGGTCACACTTCCTATTGTTCCAGGGCCTGCTATCCTGGCCTAAGTAAGATAAC 7800
7801 AGGAAATGAGTTGACTAATCGCTTATCTGGATTCTGTAAACTGACTGGCACCATAGAAG 7860
7861 AATTGATTACACATTGACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCC 7920
7921 AGGAGCCCACGCAGATGCGGACCTCCGGAGCTATTTTAAATGATTGGTCCACGGAGCGC 7980
7981 GGGCTCTCGATATTTTAAATGATTGGTCCATGGAGCGCGGGCTCTCGATATTTTAAAT 8040
8041 GATTGGTTTGTGACGCACAGGCTTTGTTGTGAACCCCATAAAAGCTGTCCCGATTCCGCA 8100
8101 CTCGGGGCCGCGAGTCCTCTACCCCTGCGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGG 8160
8161 AATAAAATCCTCTTGCTGTTTGCATCAAAAAAAAAAAAAAAAAAAAAA 8209

20/22

Figure 4.

The same nucleotide sequence as represented by bases 5260 to 8210 in Figure 3 is also representative for this Figure, with the following changes:

<u>Position</u>	<u>Change</u>
5273	G-T
5341	C-T
5351	C-T
5353	T-C
5356	C-T
5426	G-A
5464	Insertion AGA
5607	C-T
5638	C-T
5792	T-C
6191	Insertion AA
6253	T-A
6255	Insertion A
6900	C-G

Such nucleotide changes result in the following amino acid changes in the ENV polypeptide.

<u>Position</u>	<u>Change</u>
7	R-W
192	R-K
193	Deletion
194	Deletion
197	Y-Q
198	S-E
199	K-N
200	V-I
201	D-Q
204	Y-I
205	E-N
206	Insertions: G,M,S
206	L-W
208	N-I
209	S-V
211	L-Y
212	L-K
427	F-L

22 / 22

Figure 6

PPT | U3 OCT-1 c-Myb LyF-1 E47
 1 AAGAAGTGGGGAAATGAAAGGATGAAAATGCAACCTAACCCCTCCCAGAACC
 ETS-1 AP-4
 51 CAGGAAGTTAATAAAAAAGCTCTAAATGCCCCCGAATTMCAGACCCTGCTG
 NF-1 AP-1/TR
 101 GCTGCCAGTAAATAGGTAGAAGGTCACACTTCCTATTGTTCCAGGGCCTG
 ETS-1/GATA GATA ETS-1 c-Myb AP-1 GATA
 151 CTATCCTGGCCTAAGTAAGATAACAGGAAATGAGTTGACTAATCGCTTAT
 E47 AP-1
 201 CTGGATTCTGTAAAACTGACTGGCACCATAGAAGAATTGATTACACATTG
 AP-1 AP-1/GATA c-Myb AP-1
 251 ACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCCAGGAGC
 E47 ETS-1 → CCAAT
 301 CCACGCAGATGCGGACCTCCGGAGCTATTTTAAAATGATTGGTCCACGGA
 GATA → CCAAT ←
 351 GCGCGGGCTCTCGATATTTTAAAATGATTGGTCCATGGAGCGCGGGCTCT
 GATA CCAAT ← AP-1/CREB
 401 CGATATTTTAAAATGATTGGTTTGTGACGCACAGGCTTTGTTGTGAACCC
 TATA U3 | R
 451 CATAAAAGCTGTCCCGATTCCGCACTCGGGGCCGCAGTCCTCTACCCCTG
 PADS polyA
 501 CGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGGAATAAAAATCCTCTTG
 R | U5
 551 CTGTTTGCATCAAGACCGCTTCTYGTGAGTGATTGGGGTGTGCGCTCTT
 U5 | PBS
 601 CCGAKCCCGGACGAGGGGGATTGTTCTTTTACTGGCCTTTCATTGGTGC
 651 GTTGGCCGGGAAATCCTGCGACC